

(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平5-28717

(43)公開日 平成5年(1993)2月5日

(51)Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 1 B 27/024				
G 0 6 F 15/62	3 4 0	8125-5L		
H 0 4 N 5/445		7037-5C		
5/782	A	7916-5C		
		8224-5D		
			G 1 1 B 27/ 02	C
			審査請求 未請求 請求項の数1(全10頁)	最終頁に続く

(21)出願番号 特願平3-181010

(22)出願日 平成3年(1991)7月22日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 大場 章男

東京都品川区北品川6丁目7番35号 ソニー株式会社内

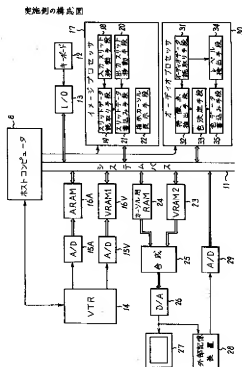
(74)代理人 弁理士 山口 邦夫 (外1名)

(54)【発明の名称】 表示装置

(57)【要約】

【目的】 音声の全体的な流れを容易に把握できる表示装置を提供する。

【構成】 VTR 14より出力されるビデオ信号およびオーディオ信号をnフレーム毎に夫々メモリ16Vおよび16Aに書き込む。イメージプロセッサ17では、サンプリング位置を水平方向に変更して垂直スリット状にメモリ16Vの画像データをサンプリングし、圧縮後メモリ23に順次書き込み、縮小画面(ビデオインデックス)を形成する。オーディオプロセッサ30では、メモリ16Aより音声データを読み取って積分処理をして音声レベルを検出すると共に、音声データの特徴点を抽出して色を決定し、メモリ23の各縮小画面の下部に位置する音声表示部に、各フレームのスリット状の画像データに夫々対応して音声レベルに応じた長さ分だけ決定した色のデータ(オーディオインデックス)を書き込む。動画と共に音声の概要を時間経過に対応させて確認でき、編集作業等が一層効率よくなる。



【特許請求の範囲】

【請求項 1】 連続する音声データの概要を順次検出する検出手段と、上記検出される概要を視覚的に表示する表示手段とを備える表示装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は、例えばビデオテープの編集に使用して好適な表示装置に関する。

【0002】

【従来の技術】 いわゆるシネフィルムのように可視画像として記録されている素材の概要の確認や所望のカットの選出は、単にその素材を目視するだけで行なうことができる。

【0003】 ビデオテープやビデオディスクなど動画像データが不可視の状態で記録されている素材の場合、その素材の概要を知るために、

(a) 1画面ずつモニタに表示して必要に応じて高速サーチなどを行なう方法

(b) モニタにマルチ画面表示によって複数のフレームの動画像を縮小してスクロール的に表示する方法（特公昭 61-4437号公報参照）等が採用されている。

【0004】 (a) の方法では、例えば 1 時間もののテレビ番組のビデオテープの概要を確認するのにそれ以上の時間がかかり、編集効率が悪くなる不都合があった。また、(b) の方法では、例えば TV コマーシャルのような短いカットを見落とすことがあると共に、その確認の再現性がなく作業によるばらつきがある不都合があった。

【0005】 そこで、本出願人は、先に一連の動画像データの概要を時間経過に対応させて圧縮した静止画像として表示し、画像の全体的な流れを精度よく認識できるものを提案した（特開平 2-260075 号公報参照）。

【0006】

【発明が解決しようとする課題】 ところで、例えばビデオテープの編集においては、画像の全体的な流れだけでなく、音声の全体的な流れを認識できればさらに便利となる。

【0007】 そこで、この発明では、音声の全体的な流れを容易に把握できる表示装置を提供するものである。

【0008】

【課題を解決するための手段】 この発明は、連続する音声データの概要を順次検出する検出手段と、検出される概要を視覚的に表示する表示手段とを備えるものである。

【0009】

【作用】 表示手段には、音声データの概要、つまり音声レベル、種類等が時間経過に対応させて連続的に表示される。そのため、音声の全体的な流れを精度よく認識で

きるようになる。

【0010】

【実施例】 以下、図面を参照しながら、この発明の一実施例について説明する。

【0011】 図 2 は、本例による動画像データの処理を示している。同図において、1 は全体としてビデオ映像群を示している。このビデオ映像群 1 は、例えばビデオテープやビデオディスク等に記録されている一連の動画像データに対応する動画像をフレーム 2 の単位で時間（t 軸）方向に並べたものと考えることができる。

【0012】 ビデオ信号のフレーム周波数は 30 Hz（NTSC 方式）であるため、ビデオ映像群 1 として 1 秒間に 30 枚のフレーム 2 が配される。2 A ~ 2 E は、ビデオ映像群 1 の一連のフレームを示している。

【0013】 4 はフレーム 2 上に設定される画像データ入力用の垂直スリット（入力スリット）であり、この垂直スリット 4 によってフレーム 2 の画像がサンプリングされる。この垂直スリット 4 は水平方向（H 方向）に所定速度で走査し、フレーム 2 の右端部に達したら再び左端部から H 方向に繰り返し走査する。したがって、垂直スリット 4 は時間軸を含めた 3 次元空間では斜めの H t 方向に走査される。

【0014】 垂直スリット 4 がビデオ映像群 1 を左端部から H t 方向に右端部まで走査するときは、f 枚のフレーム 2 を横切り、この垂直スリット 4 によって n 枚（通常は $n=1$ ）のフレームについて 1 個のスリット状の画像がサンプリングされるものと仮定する。

【0015】 f を n の倍数に選ぶと所定の整数 X を用いて、

$$f = nX$$

が成立し、f 枚のフレーム 2 からなるフレーム群（3 A、3 B 等）からそれぞれ X 個のスリット状の画像がサンプリングされる。

【0016】 本例においては、垂直スリット 4 がビデオ映像群 1 を左端部から H t 方向に右端部まで走査するのに要する時間が 12 秒に設定されると共に、 $n=1$ に設定され、

$$f = X = 12 \times 30 = 360$$

となる。

【0017】 そして、フレーム群 3 A より得られる X 個のスリット状の画像が水平方向につなぎ合わされた後に水平および垂直方向にそれぞれ圧縮されて縮小画面 6 A が形成される。縮小画面 6 A は 1 枚のフレームメモリに対応する表示画面 5 の中に嵌め込まれる。同様に、フレーム群 3 B より得られる X 個のスリット状の画像が水平方向につなぎ合わされた後に圧縮されて縮小画面 6 B が形成され、この縮小画面 6 B が表示画面 5 の中の縮小画面 6 A の隣に嵌め込まれる。

【0018】 以下のフレーム群に対応しても同様にして縮小画面が形成され、表示画面 5 中に順次嵌め込まれ

る。

【00019】実際に、垂直スリット4によってサンプリングされるスリット状の画像は1個ずつ時系列的に生成されるので、生成順に圧縮して表示画面5の中に1個ずつ嵌め込まれる。

【00020】例えば、フレーム2A～2Eに対応してそれぞれ垂直スリット4A～4EがH方向に順次走査するように割り当てられる。各垂直スリット4A～4Eでサンプリングされるスリット上の画像が、それぞれ圧縮されて表示画面5の垂直スリット（出力スリット）7A～7Eの部分の画像とされる。

【00021】垂直スリット4A～4Eでサンプリングされた画像を圧縮する方法としては、単に画像データを間引く方法や所定領域の加重平均をとる方法がある。単に画像データを間引く場合には、スリット4A～4EはH方向に1画素分の幅を有するとしてもよい。

【00022】また、表示画面5の各縮小画面の下部には音声表示部60が設けられる。各音声表示部60には、図3に拡大表示するように、各縮小画面を構成するスリット状の画像のそれぞれに対応してその画像に対応する音声レベルEaが表示される。そして、各縮小画面に対応する音声表示部60の表示は、各縮小画面に対応する音声の種類、例えば人声、音楽、その他に応じた色をもって行なわれる。

【00023】図1は、本例の表示装置を示している。同図において、8はホストコンピュータを示している。このホストコンピュータ8は装置全体の制御手段として機能する。

【00024】11はシステムバス、12はキーボードであり、オペレータはキーボード12より入出力回路13およびシステムバス11を介してホストコンピュータ8に各種コマンドを与えるようにされる。

【00025】14は動画データ源としてのVTR、15V、15AはA/D変換器、16VはビデオRAM（VRAM1）、16Aはオーディオ信号用のメモリ（ARAM）である。メモリ16Vには1フレーム分のビデオ信号が記憶され、メモリ16Aには1フレーム分のオーディオ信号が記憶される。

【00026】VTR14より出力されるビデオ信号（例えばY、R-Y、B-Y、またはR、G、Bのコンポーネント信号）は、A/D変換器15Vでデジタルデータに変換された後ビデオRAM16Vに書き込まれる。

また、VTR14より出力されるオーディオ信号はA/D変換器15Aでデジタルデータに変換された後メモリ16Aに書き込まれる。

【00027】ビデオRAM16V（VRAM1）の記憶領域は、実際の表示画面に対応して、図4に示すように水平方向（VX1方向）にHLDドット、垂直方向（VY1方向）にVLDドットとされる。このビデオRAM16Vより読み出される各画素データのアドレスは、座標

（VX1，VY1）（ $0 \leq VX1 \leq HLD-1$ ， $0 \leq VY1 \leq VLD-1$ ）で指示される。

【00028】また、17はイメージプロセッサである。ビデオインデックス作成時には、イメージプロセッサ17によって、ビデオRAM16Vの1フレーム分の画像データより垂直スリット4（図4参照）で囲まれた部分のデータが読み取られ、そのデータが圧縮されてフレームメモリよりなるビデオRAM（VRAM2）23の垂直スリット（出力スリット）7（図5参照）で囲まれた部分に書き込まれる。この他に、イメージプロセッサ17は、ビデオRAM23に対応する画面上に指示カーソルを表示する機能を有している。

【00029】イメージプロセッサ17を機能に対応した手段の集合として表現すると、このイメージプロセッサ17は入力スリット移動手段18、スリットデータ読取り手段19、出力スリット移動手段20、スリットデータ書き込み手段21および指示カーソル表示手段22よりなる。

【00030】ビデオRAM23（VRAM2）の記憶領域も、ビデオRAM16Vと同様に実際の画面上に対応して、図5に示すように水平方向（VX2方向）にHLDドット、垂直方向（VY2方向）にVLDドットとされる。このビデオRAM23に書き込む各画素データのアドレスは座標（VX2，VY2）で指示される。

【00031】また、30はオーディオプロセッサである。オーディオインデックス作成時には、オーディオプロセッサ30によって、メモリ16Aに順次書き込まれた1フレーム分の音声データが読み取られて積分処理され、音声レベルEaが検出される。そして、ビデオRAM23の垂直スリット7の下部の音声表示部60のスリット領域に音声レベルEaに対応した長さ分だけ色データが書き込まれる。

【00032】また、1縮小画面に対応する期間において、メモリ16Aよりオーディオデータが読み出され、例えばニューラルネットワークによって特徴点が抽出される。これにより人声、音楽、その他の音声の種類が判別され、その種類に応じてカラーマップより色が選択され、この色のデータが、上述したように各縮小画面の下部の音声表示部60に書き込まれる色データとされる。

【00033】オーディオプロセッサ30を機能に対応した手段の集合として表現すると、このオーディオプロセッサ30は、オーディオデータ読取り手段31、特徴点抽出手段32、色決定手段33、レベル検出手段34および色データ書き込み手段35よりなる。

【00034】24は、カーソルのデータを記憶するためのカーソル用RAMである。ビデオRAM23より読み出される画素データおよびカーソル用RAM24より読み出されるカーソルのデータは、合成回路25に供給されて合成画像データが形成される。

【00035】合成回路25より出力される合成画像デー

タは、D/A変換器26でアナログ信号に変換されてモニタ27やビデオプリンタ（図示せず）に供給されると共に、外部記憶装置（VTR、フロッピーディスク等）28にも供給される。

【0036】なお、外部記憶装置28より再生されるビデオ信号は、A/D変換器29およびシステムバス11を介してビデオRAM23に書き込みできるようにされる。

【0037】インデックス作成時において、VTR14より出力されるビデオ映像群1の画像データが、ビデオRAM16V（図4に図示）を介して一連のスリットデータとしてビデオRAM23（図5に図示）に書き込まれ、さらにオーディオデータに基づいてビデオRAM23に音声レベルEaおよび種類を示す色データが書き込まれる際の一連の動作を、図6のフローチャートに沿ってステップ毎に説明する。

【0038】この場合、ビデオRAM16Vのそれぞれのフレームから1個ずつ抽出したスリットデータをX個まとめて圧縮したものをビデオRAM23の（X×Y）個の画素よりなる縮小画像6A、6B、・・・として書き込むものとする。

【0039】[ステップ101] 次の式に従って、 ΔX 、 ΔY を計算する。

$$【0040】\Delta X = HL / X$$

$$\Delta Y = VL / Y$$

ΔX 、 ΔY は整数でなくともよく、ビデオRAM16Vの（ $\Delta X \times \Delta Y$ ）個の画素よりなるブロック33の画素データをビデオRAM23の1個の画素34の値に圧縮する。

【0041】本例においては、圧縮を簡易に行なうため、ビデオRAM16Vのブロック53の左上隅の座標（VX1, VY1）をアドレスとする画素のデータをそのままビデオRAM23の座標（VX2, VY2）をアドレスとする画素54のデータとする。 ΔX 、 ΔY が非整数の場合には座標（VX1, VY1）は整数の対ではなくなるので、座標（VX1, VY1）が指示する画素の値は周囲の画素の値からの補間によって計算する。

【0042】また、ビデオRAM23においてX個の垂直スリット7よりなる縮小画面（6A、6B等）の水平方向の配列個数hを、次の式に従って計算する。Xsは水平方向の余白の画素数である。

$$【0043】h = (HL - Xs) / X$$

また、縮小画面6A、6B・・・の番号FRをそれぞれ0, 1・・・, FR0とし、FR=0に初期設定する。

【0044】また、VTR14より出力される1フレームの画像データをビデオRAM16Vに書き込み、VTR14より対応して出力されるオーディオデータをメモリ16Aに書き込む。

【0045】[ステップ102] メモリ16Aよりオーディオデータを読み出し、オーディオプロセッサ30で、特徴点の抽出および色決定処理を開始する。

【0046】[ステップ103] ビデオRAM23の番号FRの縮小画面（6A、6B等）の左上隅の座標を（BX, BY）として、BX, BYを、次の式に従って計算する。Xs1は左端の余白の画素数、Ys1は垂直方向の上端の余白の画素数である。また、 $YA = Y + Y' + Y''$ であり、Y'は音声表示部60の垂直方向の画素数、Y''は音声表示部60と縮小画面間の垂直方向の画素数である（図5参照）。

$$【0047】BX = (FR \bmod h) \times X + Xs1$$

$$BY = [FR / h] \times Y + Ys1$$

これらの式において、（FR mod h）はFR/hの余りを示し、[FR/h]はFR/hを越えない最大の整数を示している。

【0048】[ステップ104] ビデオRAM16Vの垂直スリット4の座標VX1、ビデオRAM23の垂直スリット7の座標VX2の初期値を、それぞれ0、BXに設定する。

【0049】[ステップ105] ビデオRAM16Vの垂直スリット4の座標VY1、ビデオRAM23の垂直スリット7の座標VY2の初期値を、それぞれ0、BYに設定する。また、N=0に初期設定する。

【0050】[ステップ106] Nの値を1だけ増す。

【0051】[ステップ107, 108] イメージプロセッサ17は、ビデオRAM16Vの座標（VX1, VY1）の画素のデータを読み取ってビデオRAM23の座標（VX2, VY2）の画素のデータとして書き込んだ後に、座標VY1の値を ΔY だけ増して、座標VY2の値を1だけ増す。

【0052】[ステップ109] ビデオRAM16Vの垂直スリット4のデータをD1方向に読み出したときに、ビデオRAM23の垂直スリット7のデータはD2方向に書き込まれる。そして、ビデオRAM16Vの垂直スリット4の座標VY1がVL以下であるときにはステップ106に戻り、座標VY1がVLを越えるときにはステップ110に進む。

【0053】[ステップ110] オーディオプロセッサ30は、メモリ16Aに書き込まれている1フレーム期間のオーディオデータの全部または一部を読み出して積分する。

【0054】[ステップ111] ビデオRAM23の音声表示部60の垂直方向の画素数に対応する積分出力をDMとし、積分出力をD1とするとき、 $A = [D1 / DM \times Y']$

を計算する。 $[D1 / DM \times Y']$ はD1/DM×Y'を越えない最大の整数である。このAのデータが音声レベルEaを示すものとなる。

【0055】[ステップ112]

$DXN = VX2$

$DYN = VY2 + Y' - A$

$DYN' = VY2 + Y' - 1$

に設定する。

【0056】[ステップ113, 114] $VX1$ の値を ΔX だけ増し、 $VX2$ の値を1だけ増す。このことは、ビデオRAM16Vの垂直スリット4の位置を ΔX だけ右に移し、ビデオRAM23の垂直スリット7の位置を1だけ右に移すことを意味する。

【0057】そして、ホストコンピュータ8はVTR14より現在のフレームからn枚目のフレームの画像データおよびオーディオデータを入力して、それぞれビデオRAM16Vおよびメモリ6Aに書き込む。

【0058】[ステップ115] ビデオRAM16Vの垂直スリット4の座標 $VX1$ がHL以下であるときにはステップ105に戻る。座標 $VX1$ がHLを越えたときには、ビデオRAM16Vの垂直スリット4の水平方向への1回の走査が完了したことを意味するので、ステップ116に進む。

【0059】[ステップ116] ビデオRAM23の座標(DXN , DYN) ~ (DXN' , DYN')にオーディオプロセッサ30で決定される色の色データを書き込む。ここで、 $N=1 \sim X$ である。これにより、ビデオRAMの音声表示部60に各スリットの画像データに対応して音声レベルを示すと共に、その種類を示す色データが書き込まれることに共に。

【0060】[ステップ117, 118] 縮小画面の番号FRを1だけ増して、その番号FRがビデオRAM23の許容する縮小画面の数FR0以下であるときにはステップ102に戻る。番号FRが数FR0を越えたときには、1画面分のビデオインデックスの作成が終了したことになる。そのため、ステップ119に進んで後処理をする。

【0061】後処理としては、ビデオRAM23の画像データをD/A変換器26を介して外部記憶装置28に蓄積したり、その画像データをD/A変換器26を介してモニタ27に供給したりすることが考えられる。

【0062】外部記憶装置28には、動画像データの集合であるビデオ映像群1が一種のビデオスライスとも表現できる方法によってデータ圧縮した形で蓄積されることになる(図7に図示)。また、モニタ27には、複数の縮小画面と、各縮小画面に対応して音声レベルを示すと共に、その種類を示すインデックスが表示されることになる。

【0063】その後、オペレータの操作によって再びステップ101に戻り、VTR14からの続きのビデオ信号およびオーディオ信号に対するインデックスを作成することになる。

【0064】以上のようにして作成されるインデックスは、ビデオ映像群1のnフレーム毎に垂直スリット4の

位置を変えてサンプリングされる画像データを圧縮してつなぎ合わせた縮小画面であり、ビデオ映像群1の動画像データの概要を時間経過に対応させて確認できる。

【0065】ここで、垂直スリット4の位置を12秒で1画面の左端から右端まで走査するように変化させるため、表示画面5の中に例えばN個の縮小画像6A, 6B, ...を形成する場合は、

$12 \times N = 12N$ [秒]

$30 \times 12 \times N = 360N$ [フレーム]

より、1枚(フレーム)の表示画面5の中に12N秒分(360Nフレーム)のビデオ信号の動画像データが圧縮して表示されることになり、極めて大きな圧縮率で一連の動画像データが圧縮されている。

【0066】そして、ビデオ映像群1の動画像が12秒を基準として緩慢に変化する場合には、垂直スリット4が左端から右端へ走査しているため、略動画像の原画の状態を復元できる。一方、例えばコマーシャルのようにビデオ映像群1の動画像が急激に変化すると、縮小画面の何れかに不連続的に変化する断線が形成される。したがって、略1/360Nもの圧縮率で動画像データを圧縮して表示しているにも拘らず、緩慢に変化している部分の概要を確認することができると共に、急激に変化する部分を断線として確認できる。

【0067】また、音声表示部60には各縮小画面を構成するスリット画像に対応して音声レベルが表示される。つまり、音声レベルが時間経過に対応して表示される。また、この音声レベルの表示には各縮小画面毎に音声の種類に応じた色が付けられる。したがって、画像と共に音声の全体的な流れを容易に把握でき、編集作業等を一層効率よく行うことができる。

【0068】なお、ビデオ映像群1の個々のフレーム2より垂直スリット4を用いて画像データをサンプリングする代わりに、水平スリットを用いて画像データをサンプリングしてもよい。この場合は、水平スリットを所定速度で周期的にフレーム2の上端から垂直方向(V方向)に下端まで走査するように構成される。そして、音声表示部60は縮小画面の左または右の側部に設けられることになる。

【0069】なお、上述実施例においては、音声の種類を色を連えることで表示したものであるが、表示領域の大きさや輝度を連えることで表示するようにしてもよい。

【0070】

【発明の効果】この発明によれば、表示手段には、音声のレベルや種類等の音声データの概要が時間経過に対応させて連続的に表示されるので、音声の全体的な流れを精度よく認識できる。したがって、例えば動画像データの概要と共に使用することにより、編集作業等を一層効率よく行うことができる。

【図面の簡単な説明】

【図1】実施例の構成図である。
 【図2】インデックスの概念を示す図である。
 【図3】音声表示部を拡大表示した図である。
 【図4】ビデオRAMのデータ構造を示す図である。
 【図5】ビデオRAMのデータ構造を示す図である。
 【図6】インデックス作成時の動作を示すフローチャートである。

【図7】入力ビデオ信号とビデオインデックスとの関係を示す図である。

【符号の説明】

1 ビデオ映像群

2 フレーム

3 A, 3 B フレーム群

4 垂直スリット（入力スリット）

5 表示画面

6 A, 6 B 縮小画面

7 垂直スリット（出力スリット）

8 ホストコンピュータ

14 VTR

16 A オーディオデータ用のメモリ

16 V, 23 ビデオRAM

17 イメージプロセッサ

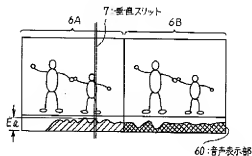
27 モニタ

30 オーディオプロセッサ

60 音声表示部

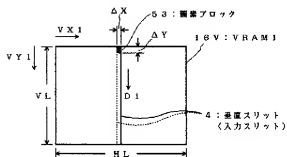
【図3】

音声表示部の拡大表示



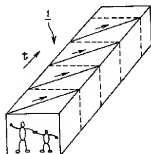
【図4】

VRAM1のデータ構造



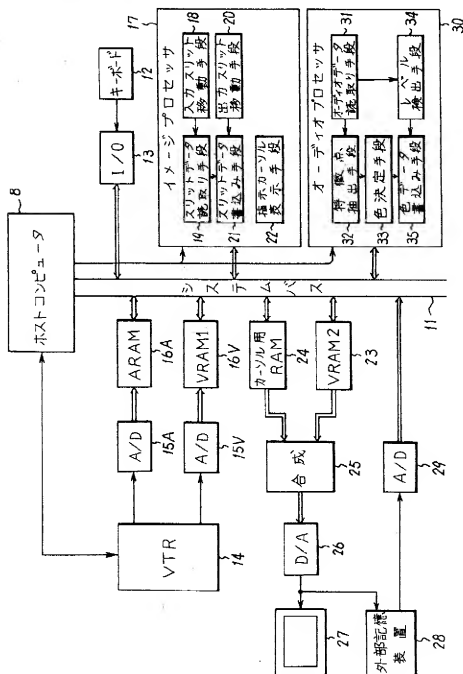
【図7】

入力ビデオ信号とビデオインデックスとの関係



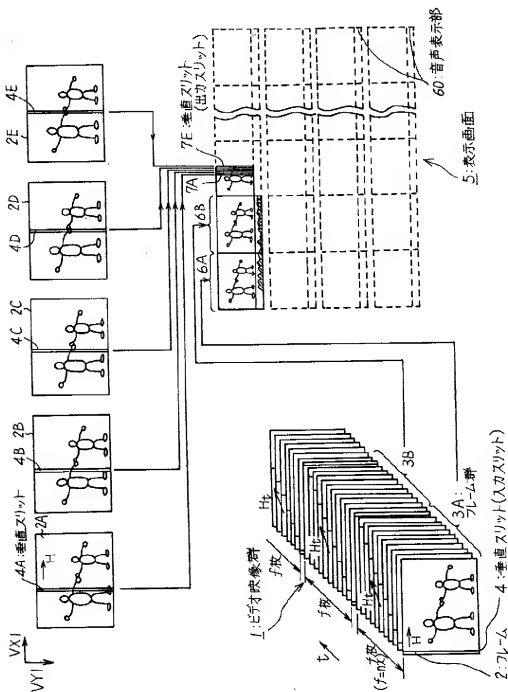
【図1】

実施例の構成図



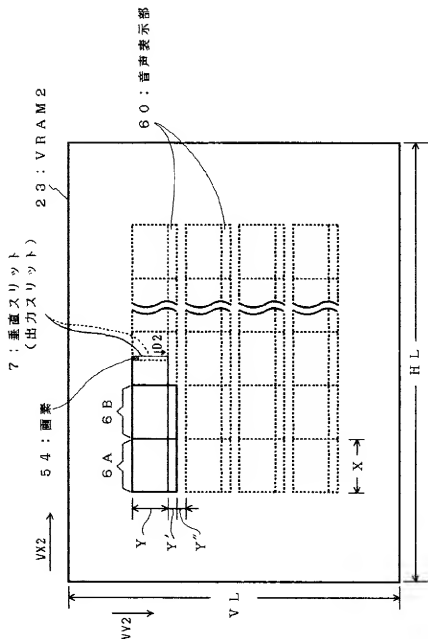
【図2】

インデックスの概念を示す図



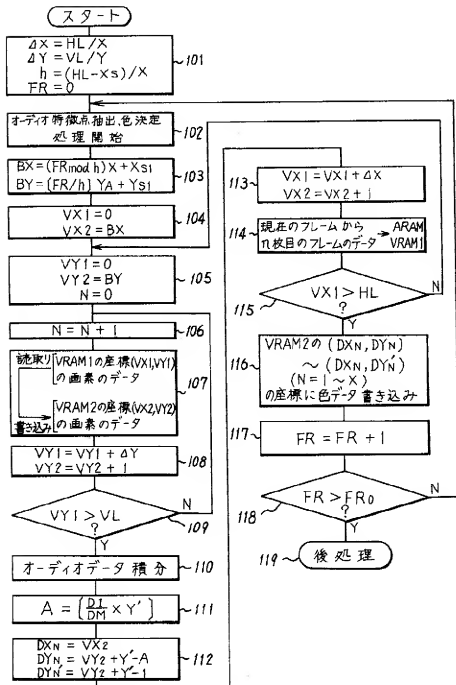
【図5】

V R A M 2 のデータ構造



【図6】

インデックスの作成時の動作



フロントページの続き

(51) Int. Cl. 5

H 04 N 5/782

5/93

識別記号

庁内整理番号

F I

技術表示箇所

C 7916-5C

C 4227-5C

G 4227-5C